

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ МЕТОДА ОБРАБОТКИ ДАННЫХ ПРИ ОГРАНИЧЕННОЙ АПРИОРНОЙ ИНФОРМАЦИИ

Марчук В.И., Саакян Г.Р., Бухтоярова И.В.

Южно-российский государственный университет экономики и сервиса
346500, Шахты, Ростовской обл., ул. Шевченко, 147, кафедра Радиотехника

В работе [1] рассматривался метод обработки данных при ограниченной априорной информации. Именно, исходные данные представлялись в виде временного ряда Y_1, Y_2, \dots, Y_n , имеющего вид последовательности дискретных отсчетов нестационарного случайного процесса $Y(t)$:

$$Y(t) = S(t) + u(t),$$

где $S(t)$ — систематическая составляющая (полезный сигнал, тренд), $u(t)$ — аддитивная шумовая составляющая, представляющая собой случайный стационарный эргодический процесс с неизвестными статистическими характеристиками. Ставится задача выделения тренда. Подобная задача может возникать: 1) в работе приемно-передающих устройств дальней или космической связи, 2) в радиотехнике при оценке помехоустойчивости схем (алгоритмов) обработки сигналов в моделируемых системах, 3) в экономических расчетах при выделении основной тенденции развития каких-либо показателей и удалении результатов воздействия случайных факторов и т.д. Функция полезного сигнала (тренда) в общем виде предполагается неизвестной (иначе задача тривиальна). К априорным данным относится предположение о том, что на некотором малом интервале времени функция $S(t)$ достаточно хорошо аппроксимируется полиномом не выше второй степени. В настоящее время решение подобной задачи осуществляется с помощью различных методов фильтрации (сглаживания), в частности, методами скользящего среднего, медианного сглаживания, экспоненциального сглаживания, моделями Бокса-Дженкинса, методами авторегрессии и ряда других [2].

В работе [1] авторами рассматривался принцип действия предлагаемого алгоритма и оценивалась его эффективность на основе имитационного моделирования. В качестве модели полезного сигнала (тренда) был использован гармонический сигнал, а в качестве модели шума — случайный процесс, подчиняющийся нормальному закону распределения.

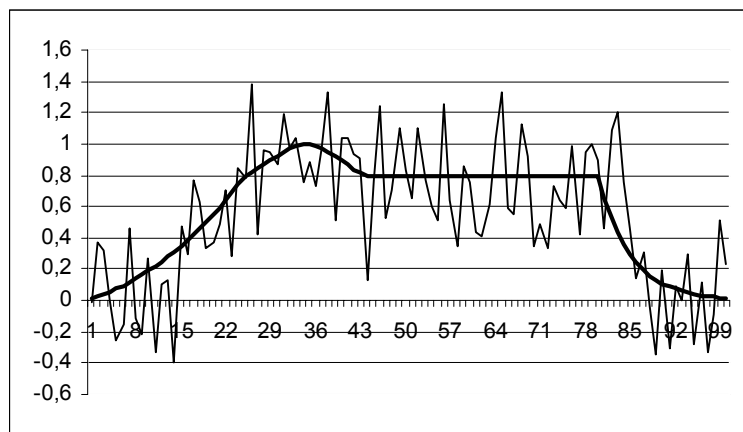


Рисунок 1.

Данная работа посвящена результатам оценки эффективности предлагаемого метода, когда функция $S(t)$ является сложной, т.е. представляет собой композицию нескольких достаточно простых функций (параболическая, синусоидальная, постоянная и экспоненциальная функции), и приводится сравнительный анализ с методом скользящего среднего. Ширина окна метода скользящего среднего выбрана равной 7, как наиболее оптимальная для обработки данного сигнала.

При этом предполагается, что исходная реализация является уникальной, т.е. она существует в единственной реализации. В качестве априорной информации о функции полезного сигнала (тренда) при использовании предлагаемого метода обработки задается средний интервал времени, на котором сигнал может быть достаточно точно описан полиномом не выше второй степени. Априорная информация о функции распределения аддитивного шума отсутствует; предполагается лишь, что аддитивная шумовая составляющая представляет собой центрированный случайный процесс. Данное предположение не является обязательным; невыполнение указанного требования приводит лишь к появлению систематической ошибки выделения тренда (на величину математического ожидания аддитивного случайного процесса).

На рисунке 1 приведена исходная реализация моделируемого сигнала с аддитивным шумом, распределенным по закону Гаусса. Обработка исходной реализации предлагаемым методом проводилась исходя из следующих предположений [1]: число разбиений исходной выборки $l=5$ (при длине выборки $N=100$, средняя длина временного интервала, на котором исходная функция полезного сигнала (тренда) приближается полиномом не выше второй степени, приблизительно равна 20); число проходов (размножений функции оценок) $m=20$; дисперсия аддитивного шума $\sigma \approx 0,3$.

На рисунке 2 представлены графики исходной функции полезного сигнала (тонкая линия) и функция выделенного предлагаемым методом полезного сигнала (толстая линия). Как видно из рисунка 2 выделенный сигнал достаточно точно описывает исходную модель сигнала при ограниченном объеме априорной информации. Как видно из рисунка 2, полученная оценка не имеет задержки во времени, характерной при обработке методами цифровой фильтрации (сглаживания). Авторами данной работы принципиально выбрано достаточно высокое значение дисперсии аддитивного шума в связи с тем, что при меньших значениях дисперсии точность выделения не позволяет визуально отличить оценку от исходного сигнала. Анализ полученных результатов обработки исходных данных позволяет сделать вывод о том, что точность выделения тренда предлагаемым методом инвариантна к виду функции исходного сигнала. Сравнительный анализ средствами имитационного моделирования показал, что предлагаемый метод при наличии вышеуказанных предположений позволяет достаточно точно описать исходную функцию (аддитивный шум отсутствует). Например, исходная функция (нормированная): гармоническая - $\sigma \approx 0,1\%$; прямоугольная - $\sigma \approx 0,9\%$; экспоненциальная - $\sigma \approx 0,04\%$; трапецевидная - $\sigma \approx 0,7\%$; сложная (рассматриваемая в данной работе) - $\sigma \approx 0,9\%$. Здесь σ - среднеквадратическая ошибка.

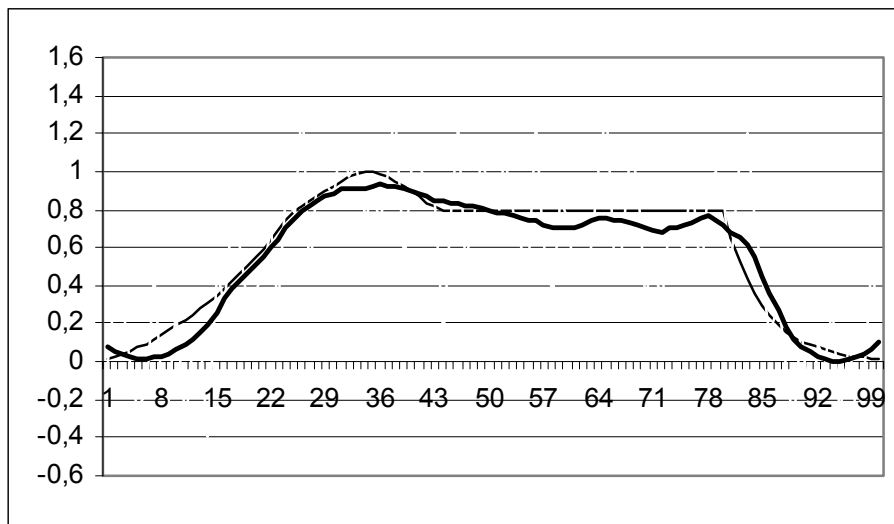


Рисунок 2

На рисунке 3 приведены результаты обработки исходной реализации методом скользящего среднего с шириной окна, равной 7. Величина ширины окна существенно влияет на точность выделения полезного сигнала (тренда) в зависимости от вида исходной функции. Уменьшение ширины окна увеличивает погрешность за счет слабой фильтрации аддитивного шума, а увеличение приводит к росту задержки сигнала по времени и, как следствие, также к увеличению погрешности выделения. Выбор оптимального значения требует либо априорной информации о функции полезного сигнала (тренда), либо эмпирического подбора. Это не позволяет рекомендовать метод скользящего среднего (и его вышеуказанные в данной работе модификации) для использования в автоматических системах обработки данных в реальном масштабе времени. В данной работе ширина окна была выбрана исходя из практических рекомендаций по применению метода скользящего среднего. Результаты имитационного моделирования показали, что среднеквадратическая погрешность выделения методом скользящего среднего отличается от среднеквадратической погрешности выделения предлагаемым методом не менее, чем в 5-10 раз (ширина окна изменялась от 5 до 9). При других значениях ширины окна результаты сравнения несопоставимо хуже. Визуальный сравнительный анализ также показывает преимущества предлагаемого метода перед известным. При этом следует отметить, что априорная информация о модели сигнала и шума практически полностью отсутствовала в алгоритмах обоих методов.

На основании вышеизложенного можно сделать следующие выводы:

- Точность выделения полезного сигнала (тренда) предлагаемым методом изменяется незначительно при изменении функции полезного сигнала (в пределах десятых долей процента).
- Предлагаемый метод обладает свойством инвариантности к функции полезного сигнала и аддитивной шумовой составляющей.

- Методами имитационного моделирования показано, что объем априорной информации, требующейся для использования предлагаемого метода сводится лишь к предположению о том, что функция полезного сигнала на некоторых интервалах хорошо описывается полиномом не выше второй степени, аддитивная шумовая составляющая является центрированной случайной функцией.

Литература

1. Марчук В.И., Саакян О.В., Саакян Г.Р. Пространственно-временная аппроксимация при выделении сигнала для случая априорной недостаточности//Изв. вузов. Сев.- Кавк. регион. Техн. науки. 1999. №3. С. 67 – 70.
2. Андерсон Т. Статистический анализ временных рядов. – М.: Мир, 1976. – 756с.

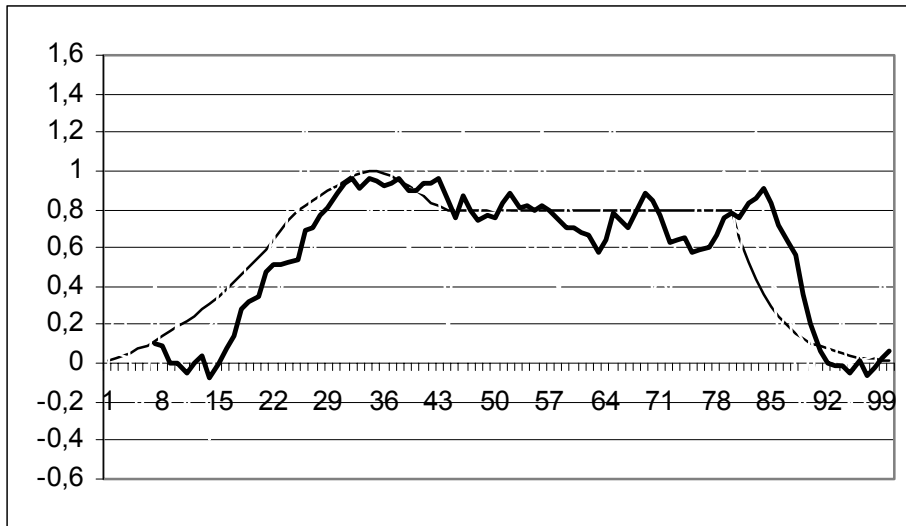


Рисунок 3

FINDING OF INVESTIGATION OF A PROCESSING TECHNIQUES OF THE DATA AT THE RESTRICTED PRIOR INFORMATION

Marchuk V.I., Saakyan G.R., Buchtoyrova I.V.

In activity [1] the processing techniques of the data was esteemed at the restricted prior information. The input data were introduced by the way of time series y_1, y_2, \dots, y_n , looking like sequence of discrete readouts of non-steady stochastic process $Y(t)$:

$$Y(t) = S(t) + u(t),$$

where $S(t)$ – systematic component (desired signal, trend), $u(t)$ – additive noise component representing random fixed ergodic process with unknowns by the statistical characteristics. The problem of a trend is put. The similar problem can arise: 1) in activity of transceiver devices distant or space communication, 2) in radio engineering at an estimation of a noise immunity of the schemes (algorithms) signal processing in simulated systems, 3) in economical calculations at allocation of the main (basic) tendency of development of any parameters (indexes) and deleting of outcomes of effect of random factors etc. The function of a desired signal (trend) in a general view is supposed obscure (differently problem is trivial). The supposition concerns to the prior data that on some small time period the function $S(t)$ is well approximated by a polynomial not above than second degree. Now solution of a similar problem implements with the help of different methods of a filtration (smoothing), in particular, methods of sliding mean, median smoothing, exponential smoothing, models of Box-Jenkins, by methods of an autoregression and number of others [2].

In activity [1] by the writers the principle of operation of tendered algorithm was esteemed and its efficiency on the basis of imitative simulation estimated. As model of desired signal (trend) the harmonic signal utilised, and as model of a noise – stochastic process subordinating to a normal distribution law.

The given activity is dedicated to outcomes of an estimation of efficiency of tendered method, when the function $S(t)$ is composite, i.e. represent composition several of enough simple functions (parabolic, sinusoidal,

constant and exponential function), and the comparative analysis with a method sliding mean is resulted. Width of a window of a method sliding mean is selected equal 7, as optimal for processing of the given signal. Thus it is supposed, that the initial implementation is unique, i.e. she exists in alone implementation. As the prior information on a function of a desired signal (trend) at usage of a tendered processing techniques the mean time period is set, on which one the signal can be precisely enough described by a polynomial not above than second degree. The prior information on a cumulative distribution function of an additive noise misses; it is supposed only, that the additive noise component represents the centered stochastic process. The given supposition is not mandatory; the default of the indicated requirement results only in occurrence (appearance) of a bias of a allocation of a trend (on value of mathematical expectation of additive stochastic process).

In figure 1 the initial implementation of a modelled signal with an additive noise distributed under the law of Gauss is adduced. The processing of initial implementation by a tendered method was conducted outgoing from the following suppositions [1]: the number of splittings of initial sampling $l=5$ (at length of sampling $N=100$, mean length of a time frame, on which one an assumed function of a desired signal (trend) comes nearer with a polynomial not above than second degree, is approximately peer 20); number of passes (breedings of a function of estimations) $m=20$; a dispersion of an additive noise $\sigma \approx 0,3$.

In figure 2 the schedules of an assumed function of a desired signal (light line) and function by a discharged tendered method of a desired signal (heavy line) are submitted. As it is visible from a figure 2, discharged signal precisely enough describes initial model of a signal at a restricted volume of the prior information. As it is visible from a figure 2, the obtained estimation has no delay in time, reference at processing by methods of a digital filtration (smoothing). The writers of the given activity in essence select high enough value of a dispersion of an additive noise that at smaller values of a dispersion the accuracy of allocation does not allow visually to distinguish an estimation from an original signal. The analysis of the obtained outcomes of processing of input data allows to draw a conclusion that the accuracy of allocation of a trend by a tendered method is invariant to a kind of function of an original signal. The comparative analysis by means of imitative simulation has shown, that tendered method if there is the above-stated suppositions allows precisely enough to describe an assumed function (additive noise misses). For example, assumed function (normalized): harmonic – $\sigma \approx 0,1\%$, rectangular – $\sigma \approx 0,9\%$, exponential – $\sigma \approx 0,04\%$, trapezoidal – $\sigma \approx 0,7\%$, composite (considered in the given activity) – $\sigma \approx 0,9\%$. Here σ is a mean square error.

In figure 3 the outcomes of processing of initial implementation by a method sliding mean with width of a window, equal 7 are adduced. The value of width of a window essentially influences accuracy of allocation of a desired signal (trend) depending on a kind of an assumed function. The decreasing of width of a window augments error at the expense of a gentle filtration of an additive noise, and the increase results in growth of a signal lag on time and, as a consequent, also to increase of error of allocation. The selection of a best value demands or prior information on a function of a desired signal (trend), or trial-and-error guard rope. It does not allow to recommend a method sliding mean (and its above-stated in the given activity of modification) for usage in automatic data-processing systems in real time. Width of a window was selected from the given activity outgoing from the practical guidelines on application of a method sliding mean. The outcomes of imitative simulation have shown that the mean square error of allocation by a method sliding mean differs from the mean square error of allocation by a tendered method not less, than at 5-10 of time (width of a window changed from 5 up to 9). At other values of width of a window outcomes of matching is non-comparable worse. The visual comparative analysis also demonstrates of advantage of a tendered method before known. Thus it is necessary to mark that the prior information on model of a signal and noise practically completely missed in algorithms of both methods.

Ground above-stated it is possible to make following conclusions:

- The accuracy of allocation of a desired signal (trend) by a tendered method changes unessentially at change of a function of a desired signal (within the limits of the tenth lobes of percent).
- The tendered method has property of invariancy to a function of a desired signal and additive noise component.
- Methods of imitative simulation have shown that the volume of the prior information requiring for usage of a tendered method is reduced only to the supposition that the function of a desired signal on some intervals is well described by a polynomial not above than second degree, additive noise component is the centered stochastic function.

The literature

1. Marchuk V.I., Saakyan G.R., Buchtoyrova I.V. Time-space approximating at a signal recovery for a case of prior failure// News High schools. North-Caucasian region. Technical sciences. 1999. N 3. P. 67 – 70.
2. Anderson T.W. The statistical analysis of time series. – M.: World, 1976. – 756 p.